

HOCRE Report No 283

CRE Accession No 28851

A Comparison of Handwriting Characteristics

Part 1

A R ALLAN, E F PEARSON, C BROWN

Home Office Central Research Establishment
Aldermaston Reading Berkshire RG7 4PN

NCJRS

APR 12 1979

ACQUISITIONS

JULY 1978

56939

SUMMARY

Simple features of prepared handwriting specimens were compared by the use of computer programs. Four handwriting specimens were taken from each of 52 individuals and no attempts were made to disguise any of these specimens. Twelve-months later two more specimens were provided, one of these in a 'normal' hand and one disguised. The similarity between a person's handwriting specimens based on eight measurements was investigated. The results of the study are presented.

INTRODUCTION

Handwriting specimens are usually compared by experts using subjective assessments of features or characteristics. For example, Angst and Erismann (1972) used a computer to help distinguish between samples and for finding authors of anonymous writing. Furthermore a statistical approach has been adopted by other workers (Kiss, 1972; Harvey and Mitchell, 1973).

At the Home Office Central Research Establishment (HOCRE) some work has been carried out on the subjective assessment of handwriting (Kind, 1978). The specimens provided for this subjective investigation were used in the work reported here as part of an objective assessment of the value of handwriting characteristics in a forensic science context. In this report we attempt to compare the specimens using simple objective measurements and comparisons.

THE HANDWRITING SPECIMENS

Three passages were used for the comparison of specimens of handwriting. These were:

- 1) *In fact, one of his converts, Ethelbert, King of Kent, who, when a heathen, had married a Christian British princess, found himself observing Easter Day on the same day that his wife was keeping Palm Sunday!*
- 2) *Under this rule, Easter Day is the Sunday following the first full moon in Spring, which is consequently known as the Paschal Full Moon; and if this full moon occurs on a Sunday, Easter Day is the Sunday following.*
- 3) *It is also of interest to note that the name of the greatest Christian Festival is derived from the name of a pagan goddess. Our Saxon forefathers kept a Spring festival in honour of the goddess Eostre.*

Fifty-two persons provided specimens and were allotted reference numbers 1-52. The specimens were referenced thus: the two copies taken of the first passage were designated specimens a and b, the second passage was designated specimen e and the third specimen f. These handwriting specimens were all written at the same time. Twelve months later, a further copy of the first passage was taken in normal hand (referenced specimen c) and another copy of the first passage in disguised hand (referenced specimen d). Thirty-eight persons provided all six specimens and twelve provided all except specimens c and d. Two persons provided only specimens c and d. Except for the disguised specimen (d) the subjects were asked to supply their handwriting specimens on A4 size lined paper in their normal hand making no attempt at disguise. For the disguised specimen the subjects were asked to disguise their handwriting in any way they thought fit. The details of the handwriting specimens taken are summarised in Table 1.

TABLE 1

SUMMARY OF HANDWRITING SPECIMEN DETAILS

Specimen	Passage	Disguise	Time (months)
a	1	NO	0
b	1	NO	0
c	1	NO	12
d	1	YES	12
e	2	NO	0
f	3	NO	0

THE INVESTIGATION

The Measurements

One measurement was taken of each eight chosen characteristics (or "dimensions") for each specimen provided by the 52 people. The measurements, which are listed in Table 2, were carried out using a transparent ruler calibrated in centimetres and millimetres. Corrections were made to dimensions 4-7 when the 11 words or spaces stretched over two or more lines.

TABLE 2

MEASUREMENTS USED FOR THE INVESTIGATION

Dimension	Measurement
1	The number of lines used in writing the specimen.
2	Margin width (on second line), left hand side.
3	Paragraph indentation of the first line (compared with the second line).
4	Length of 10 spaces between words at the end of the specimen.
5	Length of 10 spaces between words plus their 11 containing words at end of the specimen
6	Length of 10 spaces between words at the start of the specimen
7	Length of 10 spaces between words plus their 11 containing words from the start of the specimen.
8	Ratio of the average height of the full size lower case letters and the upper case letters to the height of the lower case letters (see Figure 1).

The distances measured were taken as the perpendicular intersect to the line from the end of one word to the start of the next (Figure 1(i)). In the case of (ii) this distance was taken as zero and in the case of (iii) the distance was taken as negative. When the words ran into one another as in (iv), zero distance was again taken. Example (v) in Figure 1 shows the ratio measured for dimension 8 in Table 2. A sample of the measurements is shown in Table 3.

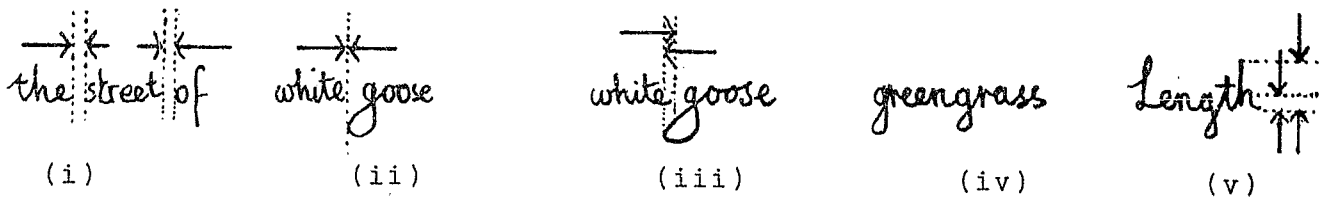
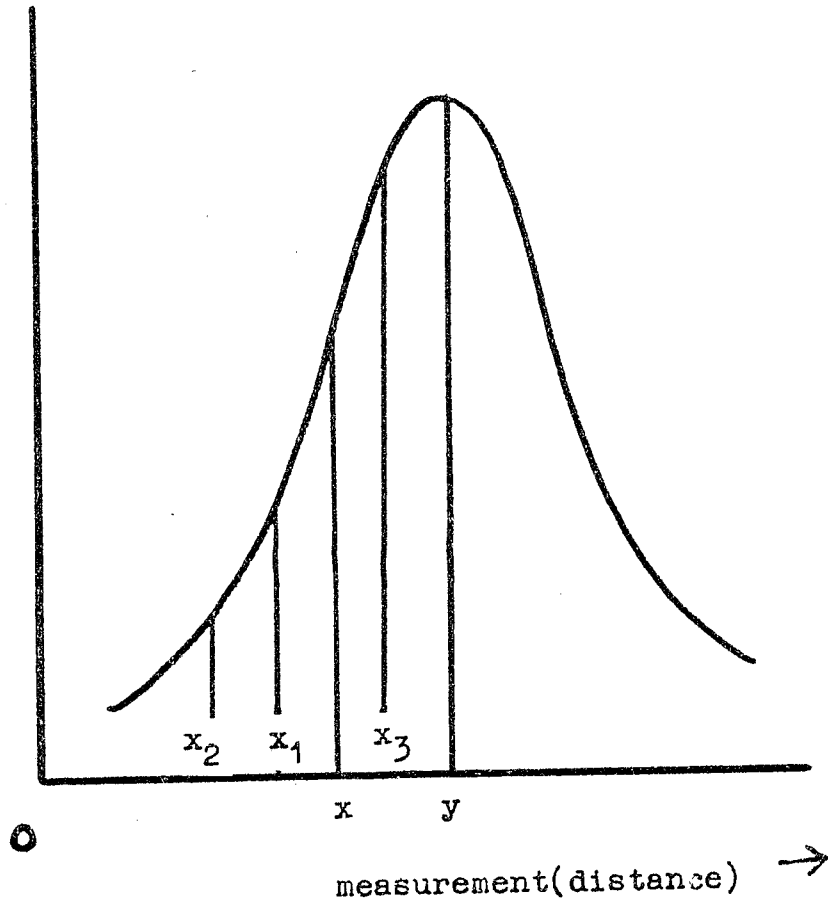


Figure 1 Examples to show how the measurements were taken

TABLE 3
SAMPLE OF THE COMPUTER PRINTOUT OF THE EIGHT
DIMENSIONS USED IN THE INVESTIGATION

Person number and specimen identification	Dimension Number							
	1	2	3	4	5	6	7	8
10 b	5.0	.5	2.0	5.6	25.1	6.6	26.4	2.4
10 c	4.5	.4	1.8	4.1	22.3	5.2	22.8	4.1
10 d	4.8	.2	1.5	4.8	20.4	5.8	20.6	2.5
10 e	4.7	.6	1.3	4.0	26.1	4.9	24.9	2.6
10 c	4.7	.7	1.4	4.7	29.4	3.8	21.3	3.0
11 a	4.4	2.1	1.3	8.2	21.7	6.2	17.3	4.2
11 b	4.2	2.1	1.5	7.3	19.4	6.3	18.3	3.6
11 c	4.4	1.5	2.1	8.7	22.0	8.0	20.7	2.5
11 d	4.6	2.0	.0	9.3	23.2	10.2	22.0	2.9
11 e	5.0	2.3	1.5	7.4	22.2	7.4	22.0	2.6
11 f	4.7	3.5	.0	7.0	24.3	6.1	16.5	3.1
12 a	3.3	.6	.8	5.2	18.5	4.7	16.1	1.9
12 b	3.1	.9	1.4	4.6	16.8	4.1	15.1	2.0
12 c	3.8	1.6	2.0	6.1	18.2	6.3	17.3	2.3
12 d	3.0	2.0	.2	4.8	14.0	5.2	14.5	2.8
12 e	3.2	.7	1.1	4.9	18.4	4.2	16.4	2.1
12 f	3.1	.7	1.0	4.9	19.0	4.6	13.7	1.9
13 a	4.8	.5	.4	5.6	26.2	6.1	24.6	2.7
13 b	4.6	.8	.1	6.1	26.6	6.1	23.7	3.0
13 c	4.8	1.1	-.3	5.8	25.7	6.2	23.6	2.6

Distribution of all measurements in the i^{th} dimension



Distance between
'crime' x and
'questioned' y
specimens = $|x-y|$;

distance between crime
and reference samples
 $= |x-x_1|$
 $= |x-x_2|$ etc
 $= |x-x_3|$

Figure 2 Diagrammatic representation of the comparison of measurements for 'crime' and 'questioned' specimens and the other reference specimens in the i^{th} dimension

The identification code combines the number allotted to an individual with the specimen number.

Calculations

The measurements were processed with a computer program run on the HP 2100A computer at HOCRE. The program compares the distances between measurements made on the samples and is called 'V2EUC'. What the program does is to calculate a distance between 'crime' and 'questioned' specimens using the measurements for these specimens in the dimensions required. This distance is compared in turn with the distances obtained from measurements on the crime sample and all the other reference specimens that have been encountered during previous experience. Figure 2 illustrates the nature of these comparisons. The process can be repeated for any combination of dimensions to determine what proportion of the reference specimens provides closer agreement to the crime specimen than the questioned specimen (Kind et al., 1978).

$$\text{Thus Distance (D) = } \sqrt{\sum_{i=1}^n \frac{(x_i - y_i)^2}{m_i^2}}$$

x_i and y_i are measurements in the i th dimension for crime and questioned specimens, and m_i is the mean value for that dimension and n is the number of dimensions. Dividing by the square of the mean m removed any possible unequal weighting between the dimensions. This comparison of multi-dimensional phenomena had been developed for other applications at HOCRE.

RESULTS AND DISCUSSION

By means of this multidimensional comparison method, all handwriting specimens of each person's writing were compared with that person's other handwriting specimens using all eight dimensions. Each of the six specimens from a person were taken in turn as the 'crime specimen' and compared with the other five as 'questioned specimens'. The file consisted of measurements from 280 specimens from 52 individuals. The object of the comparison was to ascertain the degree of correspondence or the variance in a particular person's handwriting. The computer program determined the percentage of the file which showed closer agreement to the crime specimen than the questioned specimen.

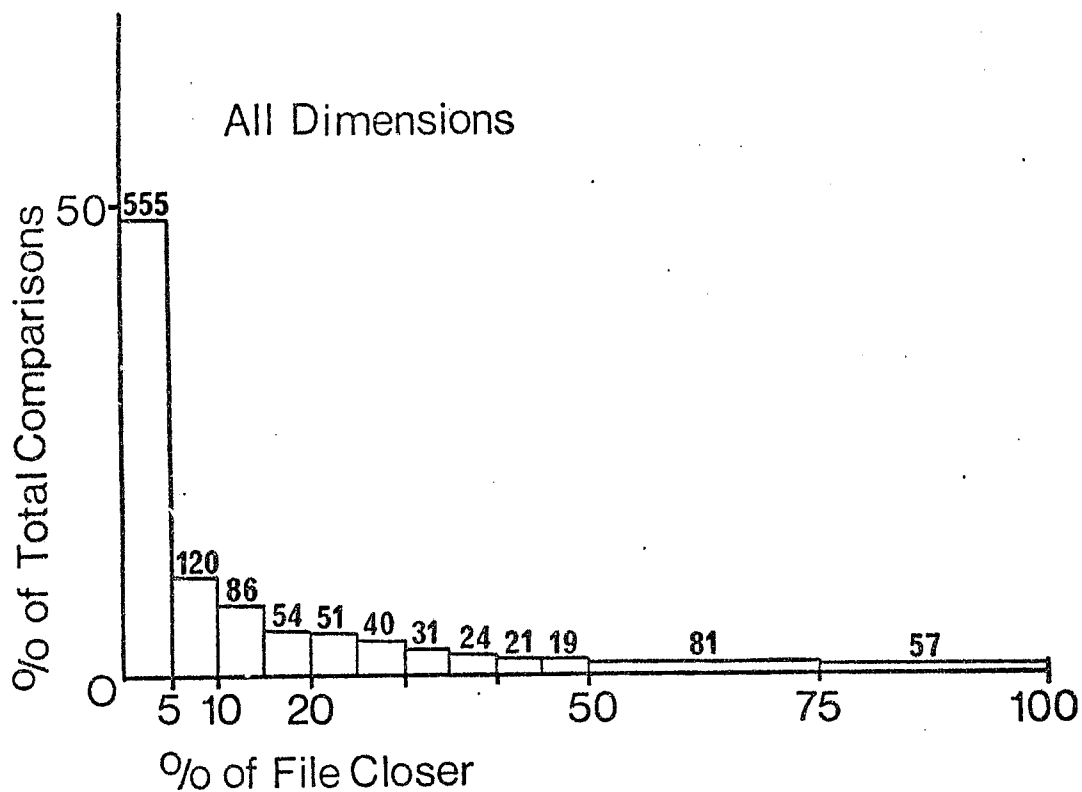


Figure 3

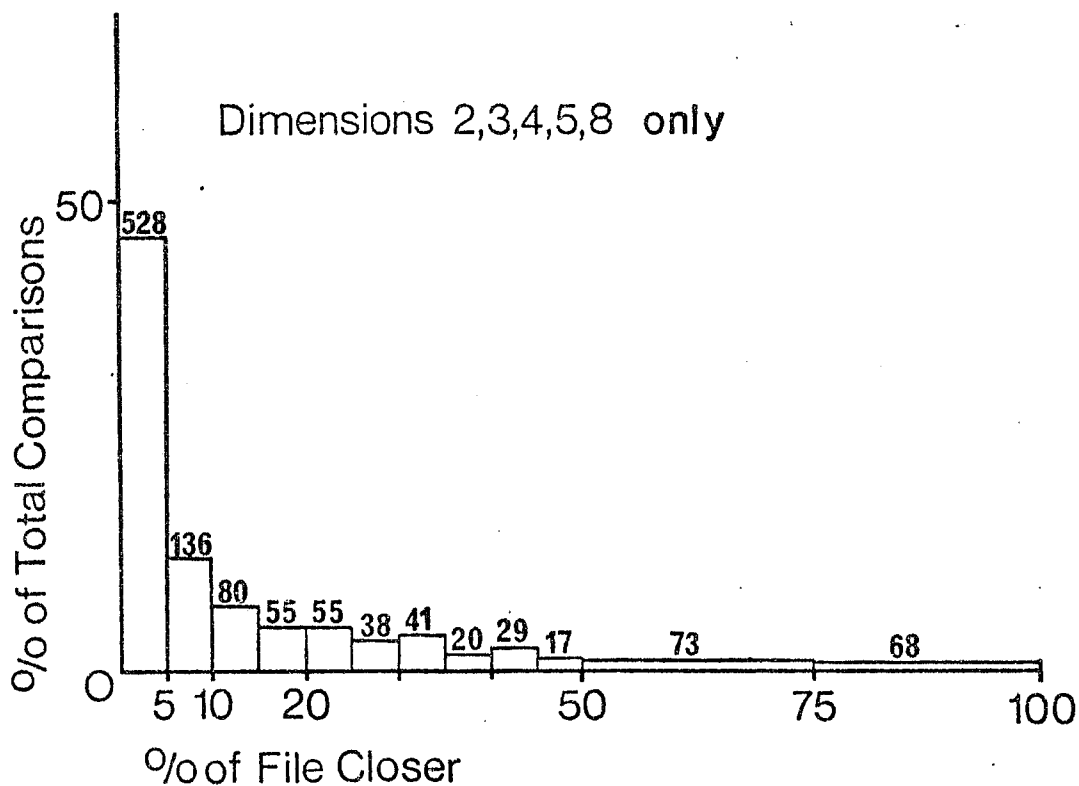


Figure 4

Figures 3 and 4

Percentages of file of specimens closer to crime than questioned specimen recovered in comparisons of all an individuals own specimens. In each case 1140 comparisons were made

The compiled results were plotted as a histogram (Figure 3). This showed that 48.7% (555 out of 1140) of the crime and questioned specimens brought out less than 5% (less than 13) of the file as being in closer agreement to the crime specimen than the questioned specimen. Of this percentage there were 8.2% (93/1140) perfect matches. However, in three cases 100% of the file was closer to the crime than the questioned specimen.

Another version of the computer program was used to produce a list of the first nine 'nearest neighbours' to the specimens from the 38 individuals who supplied all six specimens. The closest specimens are the ones with the smallest total distance (D), in all measurements, from the specimen in question. From this program two sets of results were obtained. Table 4 illustrates the number of times a person's own hand occurs in the first nine nearest neighbours, while Table 5 indicates a measure of the recoverability of the six different specimens of a person's own hand.

TABLE 4
THE NUMBER OF TIMES A PERSON'S OWN HANDWRITING
OCCURS IN THE FIRST NINE NEAREST NEIGHBOURS

Position of Nearest Neighbour (same person's writing)	Number of times specimens of same person's writing occurred in that position (including disguised specimen)
1	94
2	78
3	66
4	55
5	48
6	34
7	40
8	31
9	23
	—
Total	469

Table 4 shows that on 469 (41.1%) occasions a person's own but different handwriting specimen came within the top nine samples out of 279 different specimens. If there were to be perfect matching then a person's own handwriting would occur in the top five places on 1140 occasions (1st-5th positions).

The similarity or disparity between different specimens of the same person's writing is shown in Table 5. Specimens a and b, which are two copies of the same passage, and specimens e and f, which are copies of the other two passages show similar recoveries. (These four specimens were taken at the same time). This seems to indicate that the written content of the specimens is not particularly important in these comparisons. Specimen c which was a copy of the same passage as a and b, and specimen d which was a disguised copy of the same passage (c and d were taken twelve months after a and b) both showed poorer recovery of nearest neighbours than the other specimens. However, the effect of time, and the effect of time and disguise in a person's handwriting appear to be similar.

TABLE 5

SIMILARITY OF SAME PERSON'S HANDWRITING SPECIMENS

Specimen	Percentage recovery of same person's 5 other handwriting specimens in the 9 nearest neighbours
a	42.6
b	46.8
c	33.2
d	28.4
e	46.8
f	46.8

The above results were produced by taking all eight measurements (dimensions) into account and histograms of the distribution of all the measurements in each dimension are shown in Appendix 1. The distributions appear to be gaussian in nature.

The means and standard deviations for each dimension (assuming gaussian distribution) are shown in Table 6 and a correlation matrix for the dimensions is shown in Table 7.

TABLE 6
MEANS AND STANDARD DEVIATIONS (SD) FOR ALL THE
MEASUREMENTS (DIMENSIONS)

Dimension	Mean	SD
1	4.15	0.72
2	1.56	0.70
3	0.98	0.89
4	5.84	1.87
5	21.61	4.23
6	5.67	1.87
7	19.42	3.66
8	2.54	0.67

TABLE 7
DIMENSION CORRELATION MATRIX

Dimension	1	2	3	4	5	6	7	8
1	1.00	.12	.03	.45	.76	.39	.82	.21
2	.12	1.00	-.37	0.10	.03	-.12	.04	.03
3	.03	0.37	1.00	-.10	.03	-.12	.04	.13
4	.45	.09	-.10	1.00	.43	.77	.38	.19
5	.76	.14	.03	.43	1.00	.22	.70	.16
6	.39	.11	-.12	.77	.22	1.00	.41	.10
7	.82	-.14	.04	.38	.70	.41	1.00	.14
8	.21	.03	.13	.19	.16	.10	.14	1.00

From Table 7 it can be seen that the measurements in dimension 1 are highly correlated with the measurements in dimensions 5 and 7 and that 5 correlates with 7 also. In addition dimensions 4 and 6 correlate highly.

These correlations are not surprising since they all involve measurements of spacings in the writing of the specimens (Table 2). When the measurements in dimensions 1, 6 and 7 were omitted (Figure 4) the results obtained were very similar to those presented in Figure 3, where all the dimensions were used in the comparisons.

In order to examine the usefulness of each dimension specimens a and b were selected as duplicates and the distance between the duplicates in each dimension was computed together with the deviation from the mean value in that dimension. The quotient of the deviation from the mean divided by the duplicate distance gives a factor which

indicates how useful that particular dimension is in the similarity comparisons. It can be seen (Table 8) that dimensions 4 and 8 appear to be the least discriminating and the rest have factors up to two times better. They are all of the same order however so the significance of these differences is thus not great.

TABLE 8
DIMENSION/MEASUREMENT USEFULNESS TABLE

Dimension	Duplicate	Deviation from Mean	Factor
1	.174	.538	3.097
2	.216	.555	2.573
3	.334	.689	2.062
4	.784	1.467	1.870
5	1.124	3.129	2.785
6	.674	1.482	2.200
7	.945	1.924	3.095
8	.337	.513	1.523

CONCLUSIONS

280 specimens of handwriting were taken from 52 individuals, 40 of these providing a specimen in disguised hand. Eight sets of distance measurements were taken on these specimens and these measurements were used as a basis for the comparisons. These comparisons were performed by a computer.

Comparisons of an individual's handwriting specimens showed that 48.7% of the crime specimens brought out less than 5% of the total file as being closer than the questioned specimens (1140 comparisons). Using all dimensions in nearest neighbour comparisons a person's own handwriting specimen occupied 41.4% (469) of the first nine nearest neighbour positions including disguised specimens.

The four specimens of one person's hand which were written at the same time appeared within the first nine specimens of the file recovered with about the same frequencies. The two specimens which were written twelve months later with an attempt to disguise one of them were not recovered as well as the other four specimens. The effect of deliberate disguise was similar to the effect of time but the effectiveness of the recovery decreased by about one-third compared with the specimens written at the same time (zero time).

ACKNOWLEDGEMENTS

Our thanks to B C Platt for co-ordinating the provision of the specimens used in the assessment.

REFERENCES

Angst E, Erismann K, Kriminalistik, (1972), 2, 60

Harvey R, Mitchell R M, J. Forensic Sci. Soc., (1973),
13, 157

Kind S S, 1977, Personal Communication

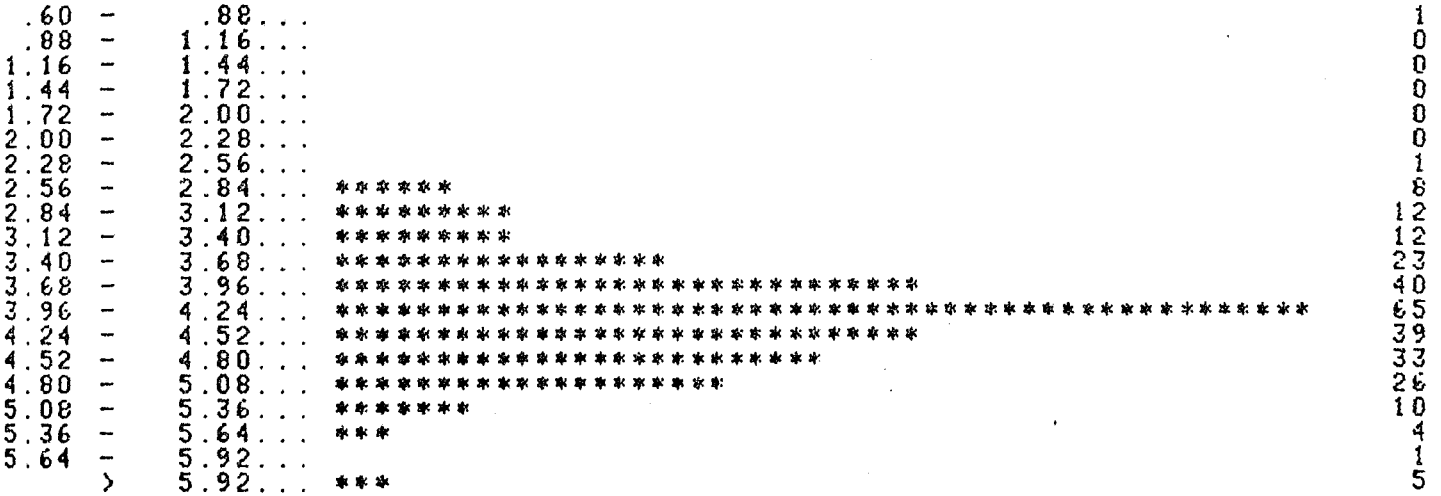
Kind S S, Smalldon K W, Brown C, Loxley D, (1978), Home
Office Central Research Establishment Report No 282

Kiss K, Krim. Forens. Wiss., (1972), 8, 73

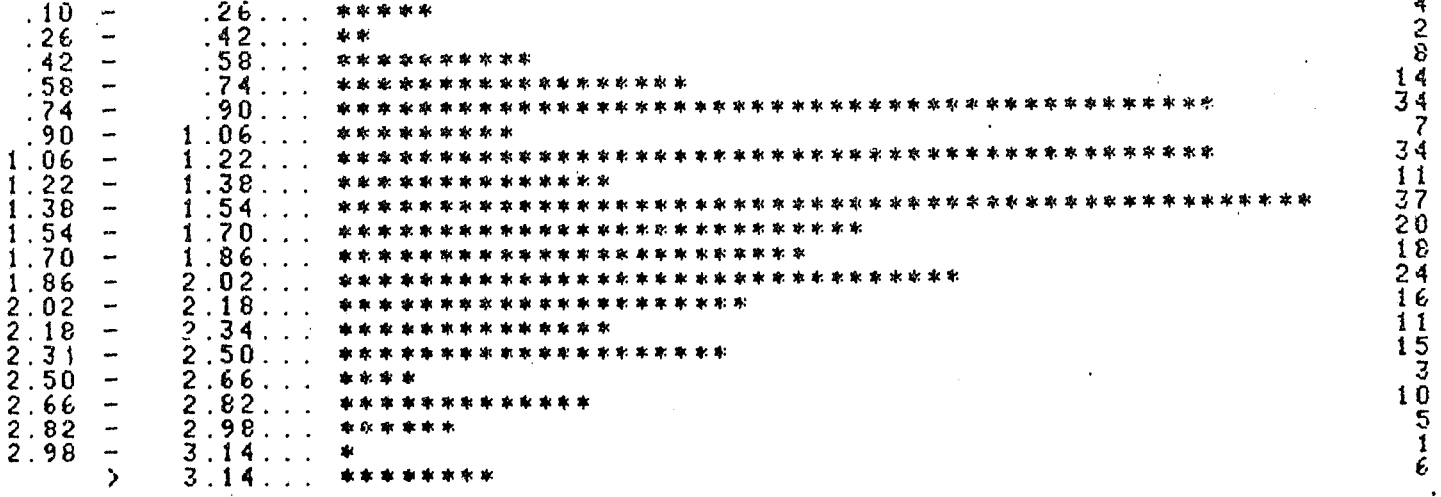
APPENDIX 1

HISTOGRAMS OF THE DISTRIBUTION OF MEASUREMENTS IN
THE 8 DIMENSIONS EXAMINED

***DISTRIBUTION FOR DIMENSION: 1**



***DISTRIBUTION FOR DIMENSION: 2**



DISTRIBUTION FOR DIMENSION: 3

- .40	-	- .22	*****	4
- .22	-	- .04	*****	10
- .04	-	- .14	*****	35
.14	-	.32	*****	40
.32	-	.50	*****	8
.50	-	.68	*****	22
.68	-	.86	*****	15
.86	-	1.04	*****	26
1.04	-	1.22	*****	26
1.22	-	1.40	*****	14
1.40	-	1.58	*****	16
1.58	-	1.76	*****	15
1.76	-	1.94	*****	15
1.94	-	2.12	*****	8
2.12	-	2.30	***	3
2.30	-	2.48	*****	7
2.48	-	2.66	*****	5
2.66	-	2.84	***	3
2.84	-	3.02	**	2
>	>	3.02	*****	6

DISTRIBUTION FOR DIMENSION: 4

2.10	-	2.52	*	1
2.52	-	2.94	*****	10
2.94	-	3.36	*****	8
3.36	-	3.78	*****	13
3.78	-	4.20	*****	23
4.20	-	4.62	*****	18
4.62	-	5.04	*****	28
5.04	-	5.46	*****	30
5.46	-	5.88	*****	24
5.88	-	6.30	*****	34
6.30	-	6.72	*****	10
6.72	-	7.14	*****	19
7.14	-	7.56	*****	13
7.56	-	7.98	*****	8
7.98	-	8.40	*****	7
8.40	-	8.82	*****	11
8.82	-	9.24	*****	2
9.24	-	9.66	*****	5
9.66	-	10.08	*	1
>	>	10.08	*****	9

DISTRIBUTION FOR DIMENSION: 5

.90	-	2.40	*	1
2.40	-	3.90		0
3.90	-	5.40		0
5.40	-	6.90		0
6.90	-	8.40		0
8.40	-	9.90		0
9.90	-	11.40		0
11.40	-	12.90		0
12.90	-	14.40	*****	5
14.40	-	15.90	*****	9
15.90	-	17.40	*****	24
17.40	-	18.90	*****	29
18.90	-	20.40	*****	40
20.40	-	21.90	*****	49
21.90	-	23.40	*****	37
23.40	-	24.90	*****	33
24.90	-	26.40	*****	22
26.40	-	27.90	*****	13
27.90	-	29.40	*****	6
>	>	29.40	*****	12

DISTRIBUTION FOR DIMENSION: 6

1.80	-	1.28	**	2
1.28	-	1.76	***	2
1.76	-	2.24	**	1
2.24	-	2.72	*	1
2.72	-	3.20	*****	11
3.20	-	3.68	*****	19
3.68	-	4.16	*****	24
4.16	-	4.64	*****	14
4.64	-	5.12	*****	44
5.12	-	5.60	*****	18
5.60	-	6.08	*****	31
6.08	-	6.56	*****	24
6.56	-	7.04	*****	24
7.04	-	7.52	*****	21
7.52	-	8.00	*****	16
8.00	-	8.48	*****	4
8.48	-	8.96	*****	6
8.96	-	9.44	*****	10
9.44	-	9.92	*****	0
	>	9.92	*****	6

DISTRIBUTION FOR DIMENSION: 7

11.30	-	12.15	***	2
12.15	-	13.00	*****	3
13.00	-	13.85	*****	8
13.85	-	14.70	*****	11
14.70	-	15.55	*****	11
15.55	-	16.40	*****	25
16.40	-	17.25	*****	29
17.25	-	18.10	*****	21
18.10	-	18.95	*****	24
18.95	-	19.80	*****	26
19.80	-	20.65	*****	16
20.65	-	21.50	*****	24
21.50	-	22.35	*****	29
22.35	-	23.20	*****	12
23.20	-	24.05	*****	14
24.05	-	24.90	*****	8
24.90	-	25.75	*****	9
25.75	-	26.60	*****	3
26.60	-	27.45	*	1
	>	27.45	*****	6

DISTRIBUTION FOR DIMENSION: 8

1.20	-	1.37	**	2
1.37	-	1.54	*****	8
1.54	-	1.71	*****	15
1.71	-	1.88	*****	8
1.88	-	2.05	*****	45
2.05	-	2.22	*****	22
2.22	-	2.39	*****	16
2.39	-	2.56	*****	45
2.56	-	2.73	*****	27
2.73	-	2.90	*****	16
2.90	-	3.07	*****	25
3.07	-	3.24	*****	15
3.24	-	3.41	*****	12
3.41	-	3.58	*****	7
3.58	-	3.75	*****	7
3.75	-	3.92	*****	4
3.92	-	4.09	*****	0
4.09	-	4.26	*****	5
4.26	-	4.43	*****	5
	>	4.43	*****	5

END